



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Filtering Algebraic Multigrid and Adaptive Strategies

A. Nagel, R. D. Falgout, G. Wittum

February 8, 2006

8th European Multigrid Conference
Scheveningen The Hague, Netherlands
September 27, 2005 through September 30, 2005

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

Filtering algebraic multigrid and adaptive strategies

Arne Nägel^{1,*}, Robert D. Falgout^{2,**}, and Gabriel Wittum¹

¹ Technische Simulation, Interdisziplinäres Zentrum für Wissenschaftliches Rechnen (IWR), Im Neuenheimer Feld 368, D-69120 Heidelberg, Germany, e-mail: {arne.naegel,gabriel.wittum}@iwr.uni-heidelberg.de

² Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Box 808, L-561, Livermore, CA 94551, USA, e-mail: rfalgout@llnl.gov

December 2005

Abstract. Solving linear systems arising from systems of partial differential equations, multigrid and multilevel methods have proven optimal complexity and efficiency properties. Due to shortcomings of geometric approaches, algebraic multigrid methods have been developed. One example is the filtering algebraic multigrid method introduced by C. Wagner. This paper proposes a variant of Wagner’s method with substantially improved robustness properties. The method is used in an adaptive, self-correcting framework and tested numerically.

1 Introduction

Geometric multigrid, as it is proposed, e.g., by Hackbusch in [4], is based on a simple observation: for problems of Poisson type usually simple iterative methods are sufficient to reduce a large portion of error components within the spectrum effectively. These methods are referred to as smoothers as the eliminated components typically belong to geometrically non-smooth eigenfunctions of the continuous operator. After several relaxation sweeps, the error is smooth in a geometric sense. The key idea of geometric multigrid is that these modes can be represented well on a coarser grid. Consequently, a coarse grid correction step is introduced in addition to the smoother.

Algebraic multigrid, as proposed by Ruge and Stüben in [5], generalizes this idea, but assumes that no information about geometry, grids and the underlying equations is available to the linear solver. The geometric notion of smoothness is substituted by the concept of algebraic smoothness. The frequently cited heuristic of classical AMG, “smooth error varies slowly in the direction of

strong couplings”, is derived from M-matrix properties and using the assumption that the constant is in the near null space of the discrete operator and needs to be interpolated almost exactly. A similar idea is used in Wagner’s method. Here, interpolation is constructed such that a certain norm of the two grid operator is minimized. Additionally, a constraint is imposed, the so called filter condition, that guarantees certain vectors to be interpolated exactly.

This paper is organized as follows: The general idea of the filtering algebraic multigrid method (FAMG) is described briefly in Section 2. In Section 3, a slightly modified version of FAMG is introduced. This approach adds robustness to the algebraic method and allows a natural embedding into classical AMG theory. In Section 4, we show the particular importance of the filter condition in the case of unknown near null space components. Following Wagner and Wittum in [8], we present an adaptive scheme which can be used to track down suitable test vectors. Section 5 concludes this work with numerical experiments underlining the theoretical results.

2 Filtering algebraic multigrid

2.1 Preliminaries

Let us first introduce some notation used in the further discussion. Our aim is to solve a linear system $Au = f$ for a non-singular $A = (a_{ij})_{i,j=1}^n$ using a multilevel method. Grid levels are indicated by a subscript $l = 0, \dots, L$, where $l = 0$ indicates the finest and $l = L$ indicates the coarsest grid. As most of the analysis is based on the two level method, we drop the index of the grid level whenever it is possible. We distinguish between coarse and fine grid nodes of a grid, forming subsets C and F respectively. For ease of presentation, we assume $A = A^T > 0$ is symmetric positive definite. Although the algorithms presented in this paper also carry over to the non-symmetric case, the main idea can be described for the symmetric case very intuitively.

* The authors are indebted to Christian Wagner for sharing his source code and contributing to this work.

** This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48

Smoothing is carried out by an iterative method with a preconditioner W . This is for instance the weighted diagonal $W = \omega^{-1} D$, $\omega > 0$, in the case of damped Jacobi smoothing or $W = (D + L)$ for Gauss-Seidel smoothing. Here we assume a decomposition $A = D + L + U$ into a diagonal part D , a strictly lower triangular part L and a strictly upper triangular part U . The error propagation operator $I - W^{-1}A$ is denoted by S . Sometimes we also use the symbol S' .

As the smoother is typically very inefficient on a small subspace of \mathbb{R}^n , we apply a coarse grid correction to approximate error components belonging to this space of the so called algebraically smooth errors. For interpolating grid functions on the fine grid from grid functions on the coarse grid, we need a prolongation P . Restriction will not be constructed explicitly, instead we use $R = P^T$, which is reasonable in the symmetric case. When a straight injection is applied, we use the injection operator R^{inj} . The coarse grid operator A_c is defined using the Galerkin operator,

$$A_c = RAP.$$

On the space of vectors, we use the Euclidean norm $\|\cdot\|_2$, and the norms $\|\cdot\|_A = \|A^{1/2} \cdot\|_2$ and $\|\cdot\|_D = \|D^{1/2} \cdot\|_2$ respectively. For matrices, we use the same symbols for the induced norms. Additionally, we consider the Frobenius norm $\|\cdot\|_F$.

2.2 Norm minimization and filter condition

We describe Wagner's method in a brief overview, omitting minor important details and refer to his original works, [6, 7], for a closer analysis. The particular choice of A_c leads to

$$(I - PA_c^{-1}RA)S = (I - PA_c^{-1}RA)(I - PR^{inj})S.$$

If we consider a submultiplicative matrix norm and assume that $\|(I - PA_c^{-1}RA)\|$ is bounded (as it is the case, e.g., for $\|\cdot\| = \|\cdot\|_A$), then finding a P that minimizes

$$\|(I - PR^{inj})S\| \quad (1)$$

yields an upper bound for the norm of the error propagation matrix of the two grid operator and influences the quality of convergence of the iterative method.

To have a notion of algebraic smoothness it is often sufficient to represent the smooth space by only one vector, namely one representing algebraic smoothness well in a local sense. Typically, this is an eigenvector associated with the smallest eigenvalue of A or an eigenvector associated with the largest eigenvalue of the smoother. We want this vector and the subspace it is representing to be interpolated (almost) exactly. Therefore, we impose a constraint

$$(I - PR^{inj})t = 0$$

or

$$(I - PR^{inj})St = 0$$

for some test vector t . For the purpose of abbreviation, we will write $Q^T = (I - PR^{inj})$. We choose to minimize (1) using the Frobenius norm, which also bounds the Euclidean norm from above. We remark that similar approaches are used in the field of sparse approximate inverse smoothers, cf. Bröker and Grote in [3]. The construction of interpolation in FAMG is based on the minimization of

$$\begin{aligned} \min_Q \quad & \|Q^T S'\|_F^2 \\ \text{s.t.} \quad & Q^T St = 0. \end{aligned} \quad (2)$$

In this constrained minimization problem S and S' may represent error propagation operators of different smoothers. If we denote the i -th column of Q (and the i -th row of $I - PR^{inj}$ respectively) by q_i , we find an equivalent local formulation for every node i :

$$\begin{aligned} \min_{q_i} \quad & \|S'^T q_i\|_2^2 \\ \text{s.t.} \quad & q_i^T St = 0. \end{aligned} \quad (3)$$

Since the prolongation should be defined locally, we fix some structural information for q_i . We use a direct interpolation, which means that coarse grid nodes are interpolated exactly, whereas fine grid nodes $i \in F$ are interpolated from parent nodes $k \in P_i$. Here, $P_i \subset C_i$ denotes a subset of coarse grid nodes $C_i = N_i \cap C$ in the neighborhood

$$N_i = \{j \neq i \mid a_{ij} \neq 0\}$$

of node i . That is, we enforce $q_i = 0$ for coarse grid nodes $i \in C$ and

$$(q_i)_k = \begin{cases} 1, & k = i \\ -p_{ik}, & k \in P_i \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

for fine grid nodes $i \in F$. Here p_{ik} is the weight we use to interpolate node $i \in F$ from node $k \in C$. For given i and n_p , we consider parent nodes $P_i \in \mathcal{P}_i$ with

$$\mathcal{P}_i = \{P \subset C_i \mid |P| \leq n_p\}.$$

After reordering the degrees of freedom, we obtain a prolongation operator of the form

$$P = \begin{pmatrix} P_{FC} \\ I \end{pmatrix},$$

where P_{FC} represents the interpolation of fine-grid values from coarse-grid ones using the entries p_{ik} from (4).

2.3 Suitable sets of parent nodes

Given a node i and a set of parent nodes $P_i \in \mathcal{P}_i$, (3) and (4) yield a quadratic problem that can be solved in a space of dimension $n_p + 1$. When the solution is unique, as it is for non-singular S' (e.g., for a convergent smoother), this process defines a mapping

$$q : \mathcal{P}_i \rightarrow \mathbb{R}^n, P_i \mapsto q_i,$$

where q_i contains the entries for the interpolation operator. At the same time, the value of the objective function in (3) can be used to compare the quality of interpolation for different sets of parent nodes. We define an admissibility condition for the parent nodes in the following way:

Definition 1. A set $P_i \subset \mathcal{P}_i$, $q_i = q(P_i)$, is a suitable set of parent nodes for node i with respect to FAMG, if

$$\theta \|S'^T q_i\|^2 \leq \|S'^T q(P)\|^2, \quad \forall P \in \mathcal{P}_i$$

and

$$\|S'^T q_i\|^2 \leq \delta.$$

The parameter $0 < \theta \leq 1$ in this definition should be chosen close to 1, say $\theta = 0.9$. It specifies that parent nodes not yielding an optimal (but an almost optimal) value of the objective function, are still acceptable for interpolation. The role of the threshold parameter δ is to prevent nodes that cannot be interpolated well enough, becoming fine grid nodes. Based on this definition, cf. Algorithm 2.2.1 in [6], Wagner proposes an iterative labeling strategy to select a coarse grid, trying to solve a trade off problem of keeping the number of coarse degrees of freedom and the number of entries in A_c as small as possible. Note that the outlined algorithm is a typical AMG setup phase: once the coarse grid is chosen and a test vector is defined, the whole process can be repeated on the next grid level until the number of grid nodes is sufficiently small.

2.4 Filter property

An interesting property that will be exploited in Section 4 is that FAMG preserves the constraint for the vector t on all grid levels. This result is independent of the objective function and only due to the constraint in (3). It can be found in [6], Proposition 3.1.1.

Proposition 1. If the test vectors on coarser levels are defined by

$$t_{l+1} = R_l^{inj} S_l t_l \quad \forall l = 0, \dots, L-1,$$

then any multigrid cycle using the pre-smoothers S_l eliminates error components in the one dimensional subspace generated by $t = t_0$ exactly.

3 Diagonally stabilized filtering algebraic multigrid

3.1 Modified objective function

Unfortunately, (2) has the property that the constrained minimization problem is sensitive to a diagonal scaling of the operator A . The same problem is known for the classical Ruge-Stüben AMG, but a fix is for instance mentioned by Brezina et al. in [2]. The idea of their argumentation will be repeated in the following subsection.

The motivation for FAMG in the previous section is based on the assumption that $\|\cdot\|_A \approx \|\cdot\|_2$ when measuring the two-grid operators of interest. To obtain a scaling-invariant method, one should use the approximation $\|\cdot\|_A \approx \|\cdot\|_D$ instead and thus try to minimize

$$\|D^{\frac{1}{2}} Q^T S' D^{-\frac{1}{2}}\|_F.$$

This yields the local problems

$$\begin{aligned} \min_{q_i} & \| (S' D^{-\frac{1}{2}})^T q_i \|_2 \\ \text{s.t.} & \quad q_i^T S t = 0 \end{aligned} \quad (5)$$

for every node i . We will refer to the multigrid method based on this quadratic problem as diagonally stabilized FAMG, or DS-FAMG. Due to the modified objective function, we alter the definition of suitable sets of parent nodes. The remainder of the setup phase remains unchanged.

Definition 2. A set $P_i \subset \mathcal{P}_i$, $q_i = q(P_i)$, is a suitable set of parent nodes for node i with respect to DS-FAMG, if

$$\theta \|(SD^{-\frac{1}{2}})^T q_i\|^2 \leq \|(SD^{-\frac{1}{2}})^T q(P)\|^2, \quad \forall P \in \mathcal{P}_i$$

and

$$a_{ii} \|(SD^{-\frac{1}{2}})^T q_i\|^2 \leq \delta$$

The choice of the objective function is still in accordance with the convergence theory developed in [6]. Yet the existence of a uniform, mesh size independent estimate

$$\|Q^T S'\|_A \leq K \|Q^T S'\|_D \quad (6)$$

with a moderate real constant $K > 0$ cannot be shown by purely algebraic means. Otherwise, the modified minimization problem leads to two grid convergence directly. Having in mind the classical definition of algebraically smooth error we note that $K = \sqrt{\lambda_{\max}(D^{-1}A)}$ yields

$$\begin{aligned} \frac{\|A^{1/2}(I - PR^{inj})Su\|}{\|A^{1/2}u\|} & \leq K \frac{\|D^{1/2}(I - PR^{inj})Su\|}{\|A^{1/2}u\|} \\ & \leq K \|D^{1/2}(I - PR^{inj})SD^{-1/2}\| \frac{\|D^{1/2}u\|}{\|A^{1/2}u\|} \end{aligned}$$

for all $u \in \mathbb{R}^n$. Although this inequality motivates our objective function, it also expresses the major difficulty we encounter. Since $\lambda_{\min}(D^{-1}A) \rightarrow 0$ for a mesh size $h \rightarrow 0$, we cannot expect the multigrid method to be convergent rapidly, as long as vectors u with $\|u\|_A \ll \|u\|_D$ are not approximately in the range of P . Yet practical experience has shown that it is often sufficient to impose the filter condition only for one representative vector, e.g. the constant for Poisson type problems.

3.2 Invariance properties

Given a symmetric positive definite matrix A and a diagonal matrix $C = \text{Diag}(\gamma_i)_{i=1}^n$ with positive diagonal

entries $\gamma_i > 0, i = 1, \dots, n$, we define a modified operator \hat{A} by

$$\hat{A} = C^{-1/2}AC^{-1/2}. \quad (7)$$

A scaling of this type can be introduced by various reasons. Obviously, \hat{A} is also symmetric positive definite and spectrally equivalent to $C^{-1}A$. Additionally, \hat{A} has the M-matrix property if and only if this is true for A .

From an algebraic point of view, scaled problems in the class of M-matrices are not more complicated than the original ones. Any robust AMG method should thus show scaling invariant convergence behavior. Numerical results, as well as the following theoretical analysis, indicate that the original FAMG does not have this property. For DS-FAMG we enforce prolongation and restriction operators to maintain some basic properties of the system matrix throughout all grids. We observe that

$$\hat{P} = C^{1/2}PC_c^{-1/2}, \quad C_c = R^{inj}C \quad (8)$$

yields

$$\hat{A}_c = \hat{P}^T \hat{A} \hat{P} = C_c^{-1/2}(P^T A P)C_c^{-1/2}.$$

If the error propagation operator of the smoother satisfies

$$\hat{S} = C^{1/2}SC^{-1/2}, \quad (9)$$

as for Jacobi, Gauss-Seidel or symmetric Gauss-Seidel with arbitrary damping or overrelaxation factors, this leads to

$$(I - \hat{P}\hat{A}_c^{-1}\hat{P}^T \hat{A})\hat{S} = C^{1/2}(I - PA_c^{-1}P^T A)SC^{-1/2}.$$

Thus the two grid operators are equivalent and we have identical two-grid convergence properties for scaled and unscaled problems. Using a recursive argument, this also carries over to the multigrid convergence.

The following lemma is the key to establish these results for the diagonally stabilized version of FAMG.

Lemma 1. *Let S, \hat{S} and S', \hat{S}' respectively satisfy (9). Additionally, assume that $\hat{K} = C^{1/2}K, \hat{t} = C^{1/2}t$ for a matrix $K > 0, t \neq 0$. Then q_i^* is a solution of*

$$\begin{aligned} \min_{q_i} \quad & \|(S'K)^T q_i\|_2 \\ \text{s.t.} \quad & q_i^T S t = 0, \\ & q_i \text{ satisfies (4)} \end{aligned}$$

if and only if

$$\hat{q}_i^* = \sqrt{\gamma_i}C^{-1/2}q_i^*$$

is a solution of

$$\begin{aligned} \min_{\hat{q}_i} \quad & \|(\hat{S}'\hat{K})^T \hat{q}_i\|_2 \\ \text{s.t.} \quad & \hat{q}_i^T \hat{S} \hat{t} = 0, \\ & \hat{q}_i \text{ satisfies (4)}. \end{aligned}$$

Proof. Let t be fixed and assume for the moment that q_i is an arbitrary vector. For

$$\hat{q}_i = \sqrt{\gamma_i}C^{-1/2}q_i$$

we find, that

$$\hat{q}_i^T \hat{S} \hat{t} = (\sqrt{\gamma_i}q_i^T C^{-1/2})(C^{1/2}SC^{-1/2})(C^{1/2}t) = \sqrt{\gamma_i}q_i^T S t.$$

Now \hat{q}_i is admissible, i.e. satisfies the constraints, for the scaled problem, if and only if q_i is admissible for the original problem. Assume q_i minimizes the original problem, but there is an admissible $\hat{p} \neq \hat{q}_i, \hat{p} = \sqrt{\gamma_i}C^{-1/2}p$, with $\|(\hat{S}'\hat{K})^T \hat{p}\| < \|(\hat{S}'\hat{K})^T \hat{q}_i\|$. As $\hat{S}'\hat{K} = C^{1/2}S'K$, division by $\sqrt{\gamma_i} > 0$ yields

$$\|(S'K)^T p\| < \|(S'K)^T q\|,$$

a contradiction. Hence, \hat{q}_i must also minimize the second problem. \square

The problems in this lemma are a reformulation of the minimization problems FAMG uses to construct restriction and prolongation. In all relevant cases S' and K are non-singular and the solution is unique. This is the case for DS-FAMG, where $K = D^{-1/2}$, and, e.g., if S' denotes the error propagation matrix of a convergent smoother.

According to this analysis, the diagonally stabilized FAMG selects points for the coarse grid independent of a diagonal scaling. Since property (8) holds for the constructed operators, we can expect identical asymptotic convergence rates.

4 Adaptive framework

Based on a priori knowledge about the discretized partial differential equation, usually some information about near null space components is known. One representative, the constant for instance, is then used as an initial test vector. If is not known a priori, several relaxation sweeps on a randomly chosen vector can be used.

Based on [8] we propose an adaptive strategy using the latest correction as test vector when the convergence rate stalls. The idea of self correction for AMG methods is also dealt with in [1] and [2].

Algorithm 1 Adaptive test vector method

```

1: function ADTV( $M_{(0)}, u^{(0)}$ )
2:   for  $i = 0$  to  $I$  do
3:      $\bar{u}^{(0)} \leftarrow u^{(i)}$ ;
4:     for  $j = 0$  to  $r - 1$  do ▷ Base iteration
5:        $\bar{u}^{(j+1)} \leftarrow \bar{u}^{(j)} + M_{(i)}^{-1}(f - A\bar{u}^{(j)})$ ;
6:     end for
7:      $t^{(i+1)} \leftarrow M_{(i)}^{-1}(f - A\bar{u}^{(r)})$ ; ▷ Adaptive part
8:      $u^{(i+1)} \leftarrow \bar{u}^{(r)} + M_{t^{(i+1)}}^{-1}(f - A\bar{u}^{(r)})$ ;
9:      $M_{(i+1)}^{-1} \leftarrow M_{(i)}^{-1} + M_{t^{(i+1)}}^{-1} - M_{t^{(i+1)}}^{-1}AM_{(i)}^{-1}$ ;
10:    end for
11:  return ( $u^{(i)}$ );
12: end function

```

Let us go through the algorithm briefly. We start iterating with a base iteration given by a preconditioner $M_{(i)}$, $i = 0$. After r iterations, the algorithm assumes that the iterative method stalls and that some components of the error cannot be eliminated. This is when the adaptive strategy is employed: we assume that the correction that is computed next, can be used as a test vector $t^{(i+1)}$. A new preconditioner $M_{t^{(i+1)}}$ satisfying

$$(I - M_{t^{(i+1)}}^{-1} A) t^{(i+1)} = 0 \quad (10)$$

is computed and one iterative step with this new method is performed. In a last step, line 9, the base iteration is updated in a multiplicative way, yielding the recursion

$$(I - M_{(i+1)}^{-1} A) = (I - M_{t^{(i+1)}}^{-1} A)(I - M_{(i)}^{-1} A)$$

for $i \geq 0$ and the respective $t^{(i+1)}$. The whole process of testing and updating the preconditioner is then repeated.

Let us first propose a modification, yielding a more flexible variant the algorithm. Instead of performing a fixed number of base iterations r in lines 4 – 6, we decide to update the preconditioner only, if we detect a “stalling” of the iterative method, which is defined in the following way: We measure the reduction rates of the residual,

$$\kappa_j = \frac{\|f - A\bar{u}^{(j)}\|}{\|f - A\bar{u}^{(j-1)}\|}, \quad j \geq 1,$$

and compute its geometric mean $\bar{\kappa}$ during the last five iterations. If this average rate is relatively large, i.e. $\bar{\kappa} > 0.5$, and only small changes occurred in the last five iterations, i.e. $|\kappa_j - \bar{\kappa}| < 0.025$, we decide, that it is appropriate to update the preconditioner and continue with line 7 of the algorithm.

According to Proposition 1, FAMG satisfies (10). However, the situation is different from the context of frequency filtering decompositions, where Wagner and Witum proposed the outlined Algorithm 1. In our case, the structure of the problem guarantees, that knowledge of one test vector is sufficient for an effective multigrid method. We thus modify line 9 in the algorithm and overwrite the preconditioner, setting

$$M_{(i+1)} \leftarrow M_{t^{(i+1)}}, \quad i \geq 0.$$

Note that this simplification would not work in a more general setting. Here the number of test vectors is crucial for the performance of the algorithm. The original version used $O(\log n)$ test vectors, which lead to a computational complexity of $O(n \log n)$.

5 Numerical Results

To give experimental justification for our considerations, we performed a variety of tests for scalar problems. Six different equations of convection-diffusion type are considered. The equations are discretized on the regularly

refined unit square, $\Omega = (0, 1)^2 \subset \mathbb{R}^2$, using linear finite volumes and a first-order upwinding scheme. The mesh width is $h = 1/512$. The boundary $\Gamma = \partial\Omega$ is decomposed into four parts, $\Gamma = \Gamma_N \cup \Gamma_E \cup \Gamma_S \cup \Gamma_W$, denoting the north, east, south and west boundary of the unit square. Different boundary conditions are applied.

The first two problems are two simple test cases.

Problem 1. Poisson’s equation

$$\begin{aligned} -\Delta u(x) &= 0, & x \in \Omega, \\ u(x) &= 0, & x \in \Gamma_N \cup \Gamma_S, \\ u(x) &= x_2(1 - x_2), & x \in \Gamma_W \cup \Gamma_E. \end{aligned}$$

Problem 2. Convection-diffusion

$$\begin{aligned} -\nabla \cdot (\nabla u(x) + v u(x)) &= 0, & x \in \Omega, \\ u(x) &= \frac{1}{2}(x_1 + x_2), & x \in \Gamma, \end{aligned}$$

with $v = (10^5, 10)^T$.

The following three problems are given by diffusion equations, each of them imposing its own characteristic difficulties for multigrid methods. Problem 3 has a discontinuous diffusion coefficient, with a square of high conductivity embedded in the interior of the domain.

Problem 3. Isolator problem

$$\begin{aligned} -\nabla \cdot (D(x) \nabla u(x)) &= 0, & x \in \Omega, \\ u(x) &= 0, & x \in \Gamma_N \cup \Gamma_S, \\ u(x) &= x_2(1 - x_2), & x \in \Gamma_W \cup \Gamma_E, \end{aligned}$$

with

$$D(x) = \begin{cases} 10^6, & 1/4 \leq x_1, x_2 \leq 3/4 \\ 1, & \text{otherwise.} \end{cases}$$

Problem 4 is an anisotropic problem, where the alignment of the anisotropy changes from the x_1 -direction to the x_2 -direction, as we approach the origin.

Problem 4. Anisotropy

$$\begin{aligned} -\nabla \cdot (D(x) \nabla u(x)) &= 0, & x \in \Omega, \\ u(x) &= x_1, & x \in \Gamma_E \cup \Gamma_W, \\ \frac{\partial u}{\partial n}(x) &= 0, & x \in \Gamma_N \cup \Gamma_S, \end{aligned}$$

with

$$D(x) = k(x) \begin{pmatrix} \epsilon & 0 \\ 0 & 1 \end{pmatrix} + (1 - k(x)) \begin{pmatrix} 1 & 0 \\ 0 & \epsilon \end{pmatrix},$$

where $k(x) = 1/(1 + e^{\alpha(x_1 + x_2 - 1)})$, $\alpha = 100$ and $\epsilon = 10^{-6}$.

Problem 5 is from the field of porous media. The diffusion tensor is introduced by a log-normally distributed scalar $D(x)$.

Problem 5. Porous media

$$\begin{aligned} -\nabla \cdot (D(x)\nabla u(x)) &= 0, & x \in \Omega, \\ u(x) &= x_1, & x \in \Gamma_E \cup \Gamma_W, \\ \frac{\partial u}{\partial n}(x) &= 0, & x \in \Gamma_N \cup \Gamma_S, \end{aligned}$$

for a log-normally distributed $D(x)$ with mean $\mu = 0$ and variance $\sigma^2 = 5.0$ for the corresponding normal distribution.

Problem 6 is a rotating convection problem.

Problem 6. Rotating convection

$$\begin{aligned} -\nabla \cdot (\epsilon \nabla u(x) + v(x)u(x)) &= 0, & x \in \Omega, \\ u(x) &= \frac{1}{2}(x_1 + x_2), & x \in \Gamma, \end{aligned}$$

with $v(x) = (-\sin(\pi x) \cos(\pi y), \cos(\pi x) \sin(\pi y))^T$ and $\epsilon = 10^{-6}$.

Restriction and prolongation are constructed by a FAMG setup phase. Afterwards, a V(1,1)-cycle is applied as a linear solver for the linear systems. The iteration used as a pre-smoother in the multigrid method is symmetric Gauss-Seidel, followed by one Jacobi step, that is only executed on the fine nodes. The iteration used for post-smoothing is the adjoint iteration. The unknowns on the finest grid are ordered lexicographically. The smoother used to construct S' in the objective function is an approximation of several damped Jacobi steps, as proposed by Wagner in [6] with $\omega = 0.85$. This routine involves a truncation of entries a_{ij} of the system operator, if

$$|a_{ij}| < \sigma \max_{k \neq i} |a_{ik}| \text{ and } |a_{ji}| < \sigma \max_{k \neq j} |a_{jk}|.$$

We use parameters $\theta = 0.95$ and $\delta = 0.1$ for the definition of suitable sets of parent nodes and interpolate from pairs of parent nodes, i.e., $n_p = 2$.

To compare the methods, we give the number of iterations m necessary to reduce the norm of the residual by 10 orders of magnitude. The iterations start with $u^{(0)} = 0$. The reduction rate of the norm of the residual in the last iteration, κ_m , is given as an estimate for the asymptotic convergence rate.

5.1 Original problems

First we compare DS-FAMG to the original FAMG for the systems generated by Problems 1 – 6. As the strength of connection within the matrix A has a physical interpretation in this case, we use threshold parameter $\sigma = 0.1$ to accelerate the setup phase. The constant vector is the test vector used for FAMG. Figure 1 indicates that both FAMG versions obtain comparable results.

Problem	FAMG		DS-FAMG	
	κ_m	m	κ_m	m
1	0.027	6	0.026	6
2	0.017	6	0.017	6
3	0.078	7	0.123	9
4	0.140	9	0.194	10
5	0.149	8	0.153	8
6	0.350	16	0.205	14

Fig. 1. Comparison of FAMG and DS-FAMG, $\sigma = 0.1$. Number of iterations m and convergence rate κ_m in the last iteration.

5.2 Scaling and normalization

In a second step, we apply a diagonal scaling as introduced in (7). For operators of this type, any notion of strength does not make sense and so we choose $\sigma = 0.0$. For reasons of comparison we evaluate the performance for the original problems,

$$C = \text{Diag}(1)_{i=1}^n,$$

first. In a next step, we normalize the systems to have a diagonal one, that is, we use

$$C = \text{Diag}(a_{ii}^{-1})_{i=1}^n.$$

Finally, we apply a random scaling of the systems using

$$C = \text{Diag}(\gamma_i)_{i=1}^n$$

with $\gamma_i = 10^{r_i}$ and uniformly distributed $r_i \in [-6, 6]$. While the constant is used as the test vector for the original systems, the constant scaled by $C^{1/2}$ is used as the test vector for the normalized and randomly scaled systems.

Figure 2 contains results obtained by Wagner's original FAMG. The three columns contain results for the three cases mentioned above. While the method performs well, or at least satisfactorily for the original and normalized systems, the results for the randomly scaled case are completely dissatisfying. Although we obtain a convergent method for problems with pure diffusion, the convergence rates are larger than the rates for the original problems. For problems containing a convective term, i.e. problems 2 and 6, the method breaks down completely.

Problem	Original		Normalized		Random	
	κ_m	m	κ_m	m	κ_m	m
1	0.026	6	0.026	6	0.852	30*
2	0.017	6	0.017	6	—	-
3	0.078	7	0.105	8	0.844	30*
4	0.133	9	0.144	9	0.801	30*
5	0.128	8	0.185	8	0.902	30*
6	0.371	18	0.346	16	—	-

Fig. 2. FAMG, $\sigma = 0.0$. Number of iterations m and convergence rate κ_m in the last iteration. The * symbol indicates that the convergence criteria were not achieved in 30 steps, and dashes indicate a diverging method.

Figure 3 illustrates a different situation for DS-FAMG. In accordance with Lemma 1, almost identical convergence rates are obtained for all problems. The number of iterations necessary to achieve the prescribed reduction of the residual is constant in most cases and differs by one in the worst case.

Problem	Original		Normalized		Random	
	κ_m	m	κ_m	m	κ_m	m
1	0.026	6	0.026	6	0.026	6
2	0.017	6	0.017	8	0.017	6
3	0.114	8	0.105	8	0.102	8
4	0.159	10	0.156	9	0.157	9
5	0.153	8	0.158	8	0.153	8
6	0.411	19	0.411	19	0.411	19

Fig. 3. DS-FAMG, $\sigma = 0.0$. Number of iterations m and convergence rate κ_m in the last iteration.

5.3 Adaptive algorithm

For the same problems, we investigate whether the adaptive strategy introduced in Section 4 is successful. The initial preconditioner

$$M_{(0)} = M_{t^{(0)}}$$

is a FAMG multigrid scheme, which is based on a test vector $t^{(0)}$. This vector is obtained from a random initial guess, that is relaxed with eight sweeps of a symmetric Gauss-Seidel iteration. The values of the random initial guess were chosen uniformly distributed in the interval $[0.5, 1]$.

Problem	Original		Normalized		Random		i
	$\kappa_m^{(i)}$	$m^{(i)}$	$\kappa_m^{(i)}$	$m^{(i)}$	$\kappa_m^{(i)}$	$m^{(i)}$	
1	0.025	6	0.025	6	0.562	8	0
2	0.012	6	0.010	6	0.022	11	1
					0.036	10	0
3	1.000	7	1.000	8	1.000	7	0
	0.104	13	0.038	12	0.109	17	1
4	0.180	10	0.191	10	0.629	10	0
					0.175	15	1
5	0.196	11	0.788	8	0.900	8	0
			0.063	14	0.110	12	1
6	0.322	15	0.361	18	0.335	17	0
							1

Fig. 4. Adaptive test vector using DS-FAMG, $\sigma = 0.0$. Number of iterations $m^{(i)}$ and convergence rate $\kappa_m^{(i)}$ in this iteration for each stage i of the algorithm from Section 4. Preliminary stages are indicated in grey ($i = 0$: initial preconditioner, $i = 1$: after an adaptive update).

For the original and the normalized problems, we discover that $M_{(0)}$ turns out to be an effective preconditioner. An adaptive update is necessary only, when the operator contains discontinuous or strongly varying coefficients, i.e. for problems 3 (original and normalized case) and 5 (normalized case only). For the randomly scaled problems, $M_{(0)}$ is only efficient for the convection dominated cases, problems 2 and 6. In all other cases, an adaptive update is required. Yet the preconditioner constructed by the adaptive test vector method turns out to be efficient and we almost obtain the original convergence properties.

6 Conclusion

The proposed DS-FAMG turns out to be a robust solver with convergence properties similar to the original FAMG. Additionally, it is robust with respect to a diagonal scaling. We showed, how the adaptive test vector method, which has been proposed in the context of frequency filtering decompositions, can be used in the context of algebraic multigrid methods.

References

1. M. Brezina, R. Falgout, S. MacLachlan, T. Manteuffel, S. F. McCormick, and J. Ruge. Adaptive smoothed aggregation (α SA). *SIAM J. Sci. Comp.*, 25(6):1896–1920, November 2004.
2. M. Brezina, R. D. Falgout, S. MacLachlan, S. F. McCormick, and J. W. Ruge. Adaptive algebraic multigrid. *SIAM J. Sci. Comp.*, to appear.
3. O. Bröker and M. J. Grote. Sparse approximate inverse smoothers for geometric and algebraic multigrid. *Applied Numerical Mathematics: Transactions of IMACS*, 41(1):61–80, April 2002.
4. W. Hackbusch. *Multi-Grid Methods and Applications*. Springer, Berlin, 1985.

5. J. W. Ruge and K. Stüben. Algebraic multigrid (AMG). In S. F. McCormick, editor, *Multigrid Methods*, volume 3 of *Frontiers in Applied Mathematics*, pages 73–130. SIAM, Philadelphia, PA, 1987.
6. C. Wagner. Ein algebraisches Mehrgitterverfahren - Entwicklung und Anwendung auf einen Sanierungsfall. Habilitationsschrift, Universität Heidelberg, 2000. (In German).
7. C. Wagner. On the algebraic construction of multilevel transfer operators (for convection-diffusion-reaction equations). In Erik Dick, Kris Rienslaagh, and Jan Vierendeels, editors, *Multigrid methods VI : proceedings of the Sixth European Multigrid Conference, held in Gent, Belgium, September 27–30, 1999*, volume 14 of *Lecture Notes in Computational Science and Engineering*, pages 264–270, New York, NY, USA, 2000. Springer-Verlag Inc.
8. C. Wagner and G. Wittum. Adaptive filtering. *Numerische Mathematik*, 78(2):305–328, December 1997.